12-2018

Effect of Real Gasoline Prices on Land Development in the United States

Markus Hildebrand
Clemson University, markus.hildebrand@outlook.com

EFFECT OF REAL GASOLINE PRICES ON LAND DEVELOPMENT IN THE UNITED STATES

A Thesis
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Master of Arts
Economics

---

by
Markus Hildebrand
December 2018

Accepted by:
Dr. Scott Templeton, Committee Chair
Dr. Babur de los Santos
Dr. Robert Fleck

ABSTRACT

Many studies have suggested that low commuting costs are one of the main factors that drive increases in the area of developed land along with income, population, and agricultural land rents. I use panel data from each state except Alaska from the years 1982 to 2012 to run a year and state fixed-effects linear regression model to determine the effect that real prices of motor gasoline have on developed land's share of non-federal land in each state. I use the log of developed land's share of non-federal land as the dependent variable and lagged values of real gasoline price, median household income, population density, and real agricultural state product per 1000 acres of rural land as variables. I find that real gasoline prices have a significant and negative effect on the share of developed land in a state. On average an increase of $1 per million BTU will result in about a 1% decrease in the share of developed land and a $1 per gallon increase will result in about an 8.7% decrease in the share of developed land.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

Land in the United States is constantly being converted from rural uses to urban uses. Land that is not considered part of the rural land base is known as developed land, and a change in the area of developed land over time is land development (NRCS, 2012). Developed land in the United States increased by 59 percent from 72 million acres to 114 million acres from 1982 to 2012 (NRCS, 2012). This includes all urban land and rural transportation land (NRCS, 2012).

Understanding the factors that drive this land development is an important part of understanding the effects of land development on other parts of society. For example, the subject of urban “sprawl” is one of much policy debate as many would like to enact policies that minimize it (Brueckner, 2000). As the economy and population grow, land must be developed for increased manufacturing, commercial, and residential use. But land development also leads to a loss in wildlife habitat and an increase in greenhouse gas emissions (Kahn, 2000, Bounoua et al, 2002). Understanding the causes of land development is important for analyzing the policies that affect land development and therefore the environment.

Theoretical and empirical studies have been done to determine the main factors that drive land development. Much of the literature points to four main variables that affect the rate at which land is developed: Income, population, agricultural land value, and commuting costs (Brueckner and Fansler, 1983, McGrath, 2005, Lubowski et al, 2008). Gasoline costs is one of the major components of commuting cost and should therefore affect the rate at which land is developed. As land gets further away from the city center, it becomes less valuable due to the increased commuting cost to get into the city. As gasoline prices and other commuting costs increase, this effect will be more pronounced, and land will lose its value more quickly as distance from the city center increases decreasing the amount of land development that takes place (McCann, 2001). I look at the theoretical foundations of this effect and use statistical methods to determine the effect empirically.

The first section discusses some of the literature that deals with the topic of land development and the economics of urban spatial structure. The second section discusses the theoretical framework for the study. The third section introduces the variables used in the empirical analysis and the data sources they were gathered from. In the fourth section, the two fixed effects regression models are developed. The results of the regressions are presented in the next section and then discussed in the last section.

CHAPTER TWO

LITERATURE REVIEW

The foundations for the study of the spatial structure of the urban economy was largely laid by Edwin Mills and Richard Muth in the 1960's and later formalized by William Wheaton (1974). Wheaton formalized the assumptions of the Mills-Muth Model and determined the sign of the effect of the specified variables on the size of cities. Among other results, he determined that as commuting cost increases, the size of a city decreases (Wheaton, 1974). He also determined that the size of a circular city would decrease with increased agricultural land rents, and the size of the city would increase with population and household income.

Several empirical studies have been done to justify the theoretical results. Brueckner and Fansler (1983) used data from 40 "urbanized areas" using explanatory variables that correspond to income, population, agricultural rents, and commuting cost and the area of the urbanized area in square miles as the dependent variable. They measured commuting costs indirectly using two different variables to represent them. Both variables relied on the assumption that increased personal vehicle usage correlated with low commuting cost. The first was the percentage of commuters using public transit, and the second was the percentage of households that have one or more vehicles. They found that the variables for income, population, and agricultural land value were all statistically significant, but both variables for commuting costs were not statistically significant.

McGrath (2005) uses more comprehensive data from metropolitan areas to build on Brueckner and Fansler. He also finds statistically significant coefficients for the variables for income, population, and agricultural rents. For commuting costs McGrath uses a regionally adjusted transportation price index, and also does not find it to be statistically significant.

More recently, empirical studies have focused on using parcel level data and new and innovative technologies for better accuracy and precision both in measuring variables and in estimating marginal effects. Several studies in the past ten to fifteen years have focused on parcel level data in and around specific regions or large cities (Chakir and Parent 2009). For example, Xiao et al (2006) analyze land use change in Shijiazhuang, China, and Liu et al (2011) have a similar study in Hangzhou. Both studies list something related to cost of transportation to the central business district as determining factors of land conversion. Many newer studies make use of new emerging technologies such as Geographic Information Systems (GIS) to analyze land use patterns (Xiao et al 2006). Some recent empirical studies also focus on the effect of economic growth in different sectors on the spatial growth of cities and the development of land. Burnett (2012) concluded that economic growth in different sectors has unique effects on the spatial growth of cities both in sign and in magnitude, suggesting that different cities will have unique spatial growth patterns under general economic growth depending on what sectors are more prevalent in that city.

CHAPTER THREE

THEORY

Demand exists for both developed land and undeveloped land. The demand for developed land is for any urban use such as commercial, manufacturing, housing, etc. The demand for rural land includes uses such as agriculture, mining, forestry, and recreation (McCann, 2001).

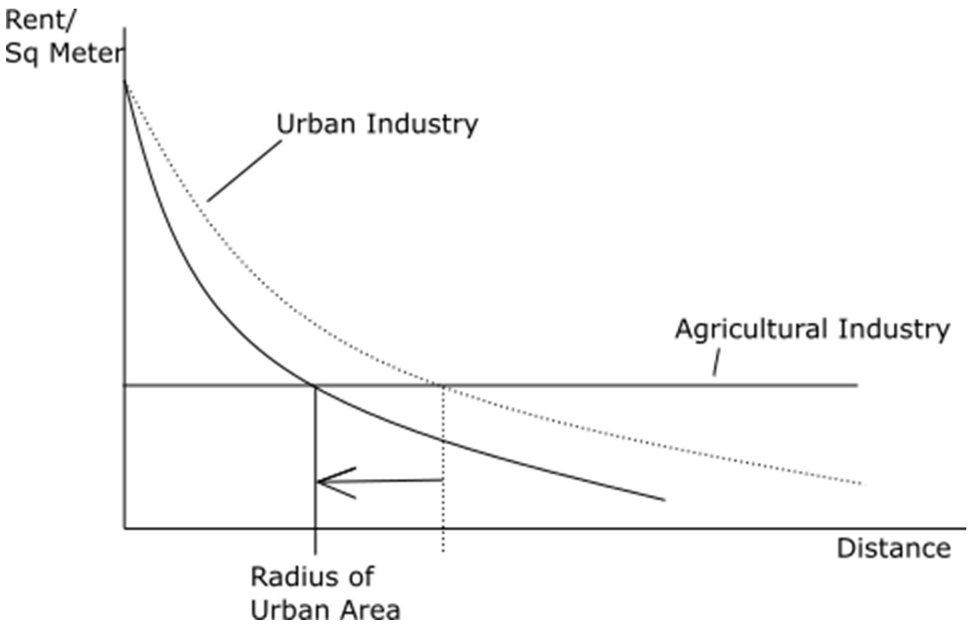
A tract of undeveloped land will be developed under two conditions. First, the tract of land will be developed if the discounted net benefits from developed use exceeds the discounted net benefits from undeveloped use. Second, discounted net income from developed land must be at least as large as the income from the undeveloped land that year plus the discounted net income from conversion a year later. If it is more profitable to wait a year and then convert the land, then the land will not be converted in the current year (Templeton and Sharma).

In this paper I am interested in how changes in the real price of gasoline affect land development. To adjust for different sizes of states, land development is defined as the change (percentage) in the total share of developed land. To understand the effect of real gasoline prices on land development, consider a single circular city with a central business district (CBD) at the center where all goods and services are traded (Wheaton 1974, McCann 2001). Businesses need to transport goods to the CBD or commute to the CBD to provide services, and consumers need to commute to the CBD to purchase goods and services (Wheaton 1974, McCann 2001). While no city truly has only one location where goods and services are traded, most cities have a downtown area with the majority of restaurants, entertainment and recreational services, retail stores, and other markets for goods and services.

As land gets further away from the CBD, it will become more expensive to commute to the CBD to trade goods and services. The increase in commuting cost as distance from the CBD increases causes the maximum rent that people who live residential dwellings are willing to pay to fall. The cost of commuting includes the real price of gasoline but also includes other things such as the opportunity cost of time. Since land is more valuable closer to the CBD, people will consume relatively less of it and live closer together resulting in higher population density. People will decide between living in small, more expensive spaces close to the CBD or larger, less expensive spaces further away from the CBD.

The effect that distance from the CBD has on willingness to pay for rent is greater for the urban sector than it is for the agricultural sector (McCann 2001). The point at which the people's willingness to pay for urban land decreases to a value less than their willingness to pay for agricultural land will be the outer edge of the urban area of the city (McCann 2001). As commuting costs increase, the willingness to pay for urban land far away from the CBD will decrease at a faster rate, which will cause the radius of the urban area of the city to decrease. Figure 1 shows the willingness to pay rent for the urban sector and agricultural sector and the effect that an increase in commuting cost per mile has on the radius of the urban area of the city (A graph similar to figure 1 can be found in McCann 2001). The radius of the city decreases simply because it is more expensive for people to commute to the CBD, so people tend to group closer together closer to the CBD.

Figure 3-1: The effect of an increase in commuting cost per mile on the radius of a circular city



For my empirical analysis in this paper, I am not looking at city-wide data, but rather state-wide data. The developed land's share of non-federal land in a state is the dependent variable of my econometric model. Non-federal land is land that has been or still could be developed. Can the theoretical analysis that I have discussed be applied on a state level the same way it is applied to a single city? When the radius of any single city in a state increases, some agricultural land must be converted to urban land. Therefore, an increase in the size of a circular city as I have discussed is equivalent to an overall increase in the developed land's share of in a state. A state is a collection of cities. Therefore, commuting cost as well as the other factors just discussed should have empirical effects on land development as they have in the theoretical models of city size. If the conditions for all cities in one state are on average more conducive to increases in urban land area than that of the cities in another state, then it stands to reason that developers in the former state would convert more land from agricultural uses to urban uses to accommodate the larger urban areas. Since a state is a collection of cities, if all or most cities are affected in a way as to increase the share of developed land, then that state will reflect that with an increase in total share of developed land.

CHAPTER FOUR

DATA AND VARIABLES

There are multiple advantages and disadvantages to using state-wide data for my analysis rather than city-wide data or data for metropolitan areas. State data allows me to account for all the non-federal land in the entirety of the United States. The coefficients will reflect the effect that the independent variables have across the entire state rather than just cities. There are no gaps between states like there are between cities used for other empirical analyses, so all the non-federal land area in the United States will be accounted for. Another advantage to using state-wide data is the availability of data for the analysis. It is much easier to find average gasoline prices for a state than it is for a city or county. I am not aware of any usable data on historical state-wide or county-wide average gasoline-prices. Gathering data such as median household income and gross state product was also much easier with states. One of the biggest disadvantages to using state-wide data is the lack of literature pertaining to state-wide land development. Most of the papers that I could find involving land development were concerned with single cities or a few cities separated by large distances.

The dependent variable for the empirical analysis is the natural logarithm of the share of non-federal land that is developed in a state. Developed land is defined as any tracts of urban land and rural transportation land (NRCS, 2012). Urban (or built-up) land is defined as "areas characterized by buildings, asphalt, concrete, suburban gardens, and a systematic street pattern (USGS LCI, 2016)."

The data for the dependent variable were gathered from the U.S. Department of Agriculture's National Resources Conservation Services. Every five years from 1982 to 2012, the NRCS has published a National Resources Inventory. Among other things, the NRI uses several methods to estimate the area of developed land in each state. These data were used to calculate the variable DEV, the developed land's share of non-federal land in each state for the years 1982, 1987, 1992, 1997, 2002, 2007, and 2012. Federal land was not considered in the analysis, because it is not readily available for development. The mean share of nonfederal developed land in a state is 9.03%. Montana in 1987 had the smallest share of developed land with 1.25% of nonfederal land being developed. New Jersey had the largest share of developed land at 36.68% of nonfederal land being developed (Table 4-1).

The independent variables in this analysis are meant to reflect the four factors that induce increases in the area of developed land according to the classic Mills-Muth model: household income, population, agricultural land rents, and commuting cost. Because, the dependent variable is only measured in five-year periods, the independent variables were averaged for the five years prior to the NRI measurements where possible. The variables were also lagged one full time period of five years. For example, the variable INCOME that affects developed land's share of non-federal land for Alabama in 2012 is the average of the median household incomes in Alabama for the years 2002 to 2006. I will discuss the reasons for lagging the independent variables later in this section.

Inflation-adjusted values for all variables were calculated using data from the U.S. Bureau of Economic Analysis. They publish tables with Gross Domestic Product by State

(GSP) for each year broken up into different industries along with Quantity Indexes for each year to convert to Real GSP. Real GSP for each state and year is measured in chained 2009 dollars. A discontinuity in the method for measuring GSP exists at 1997. Every year prior to 1997 uses the Standard Industrial Classification (SIC) industry definitions for GSP, while for years 1997 and forward the North American Industry Classification System (NAICS) industry definitions are used. This results in some inconsistency in the data and how they were gathered. The methods for estimation and sources are different, and therefore the use of all data before and after 1997 should be approached with caution, but I have chosen to combine both into one panel for this analysis. The NAICS measurements make several improvements over the SIC measurements and are therefore more accurate estimations of GSP. According to the BEA notes on this discontinuity, "Two such improvements were recognizing research and development expenditures as capital and the capitalization of entertainment, literary, and other artistic originals. These improvements have not been incorporated in the SIC-based statistics."

Along with nominal values of GSP by state and year, the BEA also publishes quantity indexes for the conversion of nominal GSP to real GSP. Just like the values for GSP, the quantity indexes are calculated using two different methods, one for pre-1997 and one for 1997 and forward. Real GSP for my model is measured in chained 2009 dollars. For all years prior to 1997, it was necessary to first convert to chained 1997 dollars. This was done by multiplying the nominal GSP by the quantity index for 1997 in 1997 dollars (100) divided by the quantity index for the given year in 1997 dollars. For example, the formula for converting GSP in 1990 from 1990 dollars to 1997 dollars is

$$GSP_{1997} = GSP_{nominal} * \left(\frac{Index_{1997}}{Index_{1990}}\right).$$

Then those were converted to 2009 dollars by multiplying by the quantity index for 2009 in 2009 dollars (100) divided by the quantity index for 1997 in 2009 dollars. For all other years the real GSP was calculated simply by multiplying the nominal GSP by the quantity index for 2009 (100) divided by the quantity index for the given year. The formula for converting nominal GSP to real GSP is

$$GSP_{real} = GSP_{nominal} * \left(\frac{Quantity\ Index_{2009}}{Quanitiy\ Index_{t}}\right).$$

The price index I used to calculate inflation-adjusted values for each state in each year was then calculated by dividing nominal GSP by Real GSP. Inflation-adjusted values for GSP, gasoline price, and income were then calculated by dividing by the price index.

The U.S. Census Bureau publishes state intercensal data tables in which are the median household incomes for each state in each year from 1984 to 2015. The variable INCOME is the previous five-year average of the state median household income converted to 2009 dollars using the state price index. The U.S. Census Bureau intercensal tables were also used to calculate the population density variable POPDENS, which is the previous five-year average population per 1000 acres of non-federal land. Population density was used rather than population to control for the vastly different sizes of states.

The metric I used for agricultural land rents is RASPAREA which is the real agricultural and mining state product per 1000 acres of undeveloped land measured in millions of 2009 dollars per 1000 acres. This was used because it reflects the productivity and therefore the rents that should be paid for the agricultural land. The BEA publishes industry-specific state products as well as industry-specific quantity indexes. These were used to calculate real agricultural state product in a similar way as the real gross state product. I included state mining product in agricultural state product, because mining also takes place on undeveloped land. The BEA considers mining a different industry, so I calculated the mining and agricultural state products separately and added them together. Wisconsin had the lowest real agricultural state product per unit rural land at $15,174.50 per 1000 acres. Rhode Island in 2012 had the highest real agricultural state product per 1000 acres of rural land at $8,004,394. Rhode Island seems to be a very high outlier for RASPAREA. If Rhode Island is ignored the maximum value of RASPAREA is $1,342,883 per 1000 acres (Table 4-1).

Finally, the U.S. Energy Information Administration publishes data on the average retail price of motor gasoline in each state in each year from 1977 to 2015. The state price index was used to convert these values to real prices. The variable GAS is the average price of motor gasoline in each state measured in 2009 dollars per million British thermal units averaged over the five years prior to the NRI measurements in the dependent variable. There are about 115,000 BTU in one gallon of gasoline, so a price of $20 per million BTU is about $2.30 per gallon. Gasoline prices range from $8.78 per million BTU to $30.51 per million BTU with a mean of $16.45 per million BTU. That is equivalent to about $1.01 per gallon (Michigan, 2002) to about $3.51 per gallon (Nevada, 1987) with a mean of about $1.89 per gallon in 2009 dollars.

As I mentioned in the beginning of this section, all the explanatory variables were lagged one full five-year time-period. This was done because land development takes time to complete and to avoid simultaneity problems. Land development does not usually take place exactly when the decision is made to develop the land. It takes a few months to years to plan and make the conversion. Lagging the variable one time-period helps control for this by looking at the conditions the state was in five years prior to conversion, which might better estimate the conditions when the decision to develop was made. It also helps to eliminate simultaneity problems by taking independent variable measurements before the dependent variable. This helps to establish causality and reduces the effect of reverse causality of the dependent variable on the independent variable. I lagged the variables a full five years to avoid any overlap between the period of independent variable measurements and dependent variable measurements. As an example of how this could cause issues, consider if the independent variable INCOME were only lagged one year so that for the year 2012, the variable INCOME were the average median household income from 2007-2011. Theory says that the share of developed land should increase with higher median household income, but if we measure the variables this way it could be possible that there is average to above average income in 2008 and above average land development with it. But then in 2011, there could be income well below the average bringing the overall value for the variable INCOME down below average, but there would still be above average land development due to the high incomes in 2007-2008. In this case we would be attributing the increase in land development in 2008 to a below average income level in 2011, which does not make sense. To determine the effect of the price of gasoline on the decision of land developers to develop this land, we do not want to look at the gasoline prices for the years the land development took place, but rather the five years prior to the land development taking place. The variable GASPRICE is lagged one time period so that for the development that took place between the years 1987-1991, we are looking at the gasoline prices from 1982-1987.

Table 4-1: Descriptive Statistics for all variables

Variable	Obs.	Mean	Standard Deviation	Minimum	Maximum
Share of Developed Land	294	0.0903	0.0757	0.0125	0.3668
Real Agricultural State Product per 1000 Acres of rural land (Chained 2009 dollars per 1000 acres of rural land)	293	324992	639801	15175	8004394
Population Density (People per 1000 acres of nonfederal land)	294	283	344	12	1733
Median Houshold Income (Chained 2009 dollars)	294	47676	9261	25341	81363
Gasoline Price (Chained 2009 dollars per million BTU)	294	16.45	4.30	8.78	30.51

CHAPTER FIVE

EMPIRICAL SPECIFICATION

The model used for the analysis is a fixed effects log-level linear regression model. The model takes the following form:

$$\ln(devshare) = \alpha_{it} + \beta X_{it} + \epsilon_{it}$$

This model uses both state fixed effects and years fixed effects, so α_{it} is a state, $i = 1, 2, 3, \dots$, and year, $t = 1, 2, 3, \dots$, specific intercept term that encompasses and controls for unobserved variation between state and years that could be correlated with the explanatory variables. X_{it} is the 1 x 4 row vector of explanatory variables in state i in time t, and β is a 4 x 1 column parameter vector.

The equation $\ln(devshare) = \alpha_{it} + \beta X_{it}$ is a linear estimation of the natural log of the share of developed land in a state with the explanatory variables we have discussed. The error term ϵ_{it} represents random processes and any variables omitted due to researcher ignorance, and the mean is assumed to be zero.

The model was also fitted with squared terms for each of the explanatory variables. The variables RASPAREASQ, INCOMESQ, POPDENSQ, and GASPRICESQ are the squares of RASPAREA, INCOME, POPDEN and GASPRICE. This was done to better fit the data and test for any turning points where the effect of a given variable could change.

The regression was performed using Stata. The cross-sectional and time variables were set using the xtset command, and then the squared terms were generated. The xtreg command was used to perform a fixed effects regression. Xtreg only performs fixed effects by the cross-sectional categories by default, so the time fixed effects were added manually, by adding a dummy variable for each of the years 1992, 1997, 2002, 2007, and 2012. The base year was 1987.

CHAPTER SIX

RESULTS

The coefficients, standard deviations, t-statistics, and p-values for the regression can be found in table 2. Because squared values for each variable are included in the regression, the coefficient does not represent the marginal effect of the variable on the log of the share of developed land. The marginal effect of a variable can be found by taking the partial derivative of the variable with respect to the dependent variable. For example, the marginal effect of gasoline price on the natural log of the share of developed land is $\frac{\partial \ln(devshare)}{\partial gas}$. The model is of the form

$$\ln(devshare) = \alpha + \beta_{gas}gas + \beta_{gassq}gas^2 + \gamma X_{it} + \epsilon_{it}$$

where X_{it} is a vector containing all the other explanatory variables other than GAS and GASSQ. The marginal effect is then the partial derivative with respect to GAS, which is

$$\frac{\partial \ln(devshare)}{\partial gas} = \beta_{gas} + 2\beta_{gassq}gas$$

The marginal effects of each variable at the mean value of the variable are given in table 3. These marginal effects are log-level effects meaning that they reflect a percentage change in the dependent variable. For example, at the mean gasoline price, a \$1 per million BTU increase in the price of gasoline results in a 0.93% decrease in the share of developed land. It is important to note that this is not a decrease of 0.93 percentage points but a decrease of 0.93% of whatever the share of developed land was before. So, if the share of developed land is at 0.10, and the average gasoline price rises by \$1 per million BTU then we expect the share of developed land to increase by $0.10 * 0.00932 = 0.000932$ or 0.0932 percentage points, so the new share of developed land is 0.100932.

Table 6-1: Estimated Model of Natural Logarithm of Share of Developed Land: Quadratic Specification

Variable	Coefficient	Standard Deviation	t	p
rasparea	0.065419	-0.0344226	1.9	0.059
raspareasq	-0.0140513	-0.0068697	-2.05	0.042
popdens	0.0011153	-0.0002836	3.93	0
popdenssq	-0.000000351	-0.000000131	-2.68	0.008
income	0.0000145	-0.00000549	2.64	0.009
incomesq	-1.7E-10	-4.61E-11	-3.67	0
gasprice	-0.0210994	-0.008958	-2.36	0.02
gaspricesq	0.0003581	-0.0002051	1.75	0.083

Table 6-2: Average Marginal Effects of Gas Prices and Three Other Variables on the Natural Logarithm of Developed Land's Share of Non-federal Land

Variable	Average Marginal Effect
Rasparea	0.05629
Popdens	9.1663E-04
Income	-1.7099E-06
Gasprice	-0.00932

The overall R-squared is 0.7285. The F-statistic for all coefficients is 141.97, which is large enough to conclude the regression as a whole is statistically significant. All coefficients are statistically significant at a 10% significance level. All coefficients except rasparea and gassq are significant at a 5% significance level.

Because the variables have squared terms, it is also necessary to find the point at which the sign of the marginal effect of each variable changes. This can be done by setting the partial derivative with respect to each variable equal to zero and solving for the variable. This gives us the point at which the sign of the marginal effect changes as

$$-\frac{\beta_x}{2\beta_{x^2}}$$

for each variable x. The marginal effect of RASPAREA is positive up to a value of \$2.33 million per 1000 acres of rural and then becomes negative. The marginal effect of POPDENS is positive up to a value of 1,588 people per 1000 acres. The marginal effect of INCOME is positive up to a value of \$42,657 per household. The marginal effect of GASPRICE is negative up to a value of \$3.44 per gallon.

CHAPTER SEVEN

DISCUSSION

The model shows a negative relationship between gasoline prices and land development, supporting the theoretical prediction. The squared term for real gasoline prices is not significant at the 95% significance level but is at the 90% significance level. If we do account for the quadratic term, the effect is still negative up to a value of about \$29.46 per million BTU or \$3.44 per gallon. This is at the very upper end of the range of observed gas prices. The average marginal effect is -0.00932. So, for each \$1 per million BTU increase in real gasoline price, we expect about a 1 percent decrease in the share of developed land in a state. 1 percent might seem like a small amount of change, but it is important to note that the range of real gasoline prices range from about \$9 per million BTU to about \$30 per million BTU, so a change of \$10 is not unrealistic and could result in about a 10% change in the share of developed land, which is far from insignificant. Converting to real price per gallon, a \$1 per gallon increase is the same as a \$8.70 per million BTU increase and, using the average marginal effect, would result in about an 8.7% decrease in the share of developed land. On average if the share of developed land in a state decreases by 8.7%, that would be a 0.79 percentage point decrease. The data clearly show that the real price of gasoline in a state has a significant impact on how much land is developed.

Real gasoline prices are not the only measure of commuting cost. Other costs should be considered in future research. The most important cost that was not accounted for in this analysis is opportunity cost of time spent commuting. This would be hard to account for in an empirical study like this one, but average wages in a state would be a good place to start. Wages can be used to estimate how much people's time is worth to them and can be a good measure for opportunity cost of time. Another thing to consider is traffic flow within cities. This is closely related to opportunity cost of time, but in areas with very bad traffic flow, people are much more likely to not want to commute from far away into the city because of the cost it incurs.

The coefficient for population density is positive for all but the absolute highest values. This is consistent with theory. As more people move into a state, more land will be developed to meet the increased demand for urban housing. It makes sense that in very populated states such as Massachusetts, the effect of population will diminish, because there are already so many people there. It is possible that this reflects populated cities' tendency to build upward rather than outward with many tall apartment buildings and skyscrapers.

The marginal effect of RASPAREA is positive for most of the range of the data. Theory suggests that as the rent for agricultural land increases, it is less likely to be developed, because high land rents correlate to relatively higher returns for agricultural use compare to urban use. According to this theory, we should expect a negative marginal effect for these variables, but the analysis shows a positive effect. The marginal effect of RASPAREA is positive for almost all values in the range of the data.

There could be a few reasons for this discrepancy between theory and empirical results. My proxy for land rents reflect the average productivity of rural land across the whole state when what matter most are the rents of rural land adjacent to urban areas. This could be a good reason not to expect the same results for state-wide analysis as we would expect for city-wide or metropolitan area analysis. Other studies mentioned have used data from specific metropolitan areas and have used more local farmland rents.

The variable for real agricultural and mining state product per unit area of undeveloped land could be affected by omitted variable bias. It is possible that the real agricultural and mining state product per unit area of undeveloped land is correlated with real non-agricultural and non-mining state product per unit area of developed land. If these two variables are positively correlated, it would explain the positive coefficient. In future research, I would like to add another variable to control for non-agricultural and non-mining state product per unit area of developed land.

The marginal effects for INCOME are also inconsistent with theory. We would expect the marginal effect of median household income in a state to increase the amount of land development, but we find from the empirical analysis that the effect is only positive for values lower than around $42,000, which is slightly less than the mean for the data set. When median income of a state rises above the mean of medians for the country, the data show that less land is developed. A few hypotheses could help explain this. Firstly, like the issue with agricultural land value, the state-wide data include incomes of people living far away from built up areas, and this could cause discrepancies between the analysis done here and city-wide studies. It is possible that as income increases past the mean, that more people choose to live farther away from the city. People who can afford the commuting costs and large tracts of rural land might prefer to live there rather than in built up urban areas. This effect might be becoming more pronounced since the studies done in the 80s to early 2000s, because technological advancement allows for easier communication from rural areas. One of the main assumptions of our model was that people must commute to the central business district to carry out business, but this is becoming less and less true with the advent of the internet and smart phones. Many people can afford to live in rural areas and stay connected and work from home. I am not aware of any research that suggests this is happening though. There is growing evidence that suggests that more people are moving away from big cities to suburbs, however, but suburbs are considered developed land (Kolko, 2017).

CHAPTER EIGHT

CONCLUSION

My research shows that the real price of motor gasoline in a state has a negative effect on land development. The coefficient estimates in the state and year fixed-effects log linear regression model were statistically significant and consistent with theory that suggests a decrease in urban spatial expansion with an increase in real gasoline price. My research shows that a $1/gallon increase in the real price of gasoline results in an average of 8.7% decrease in developed land holding all other variables constant.

This study should help policy-makers better understand how gasoline prices effect land development in the United States by confirming that low gasoline prices will encourage people to live further apart. Many people are interested in the spatial structure of cities for various reasons. Many studies have been done to assess the impact of urban sprawl on the environment and the social structure of cities. Knowing that low gasoline prices encourage outward growth of cities helps us understand one of the key factors that could be causing this sprawl. Many policies try to fight urban sprawl by introducing strict zoning laws. It is important that policy makers understand the market forces that determine what land is developed, so that they can better assess if these market forces should be suppressed or if they should allow the market to allocate land free from regulation. If the goal of a policy-maker is to reduce sprawl, the results of this paper could even be a reason to introduce higher taxes on gasoline to drive up prices to encourage people to live closer together and not expand outward.

My research finds that the effect of population is significant and consistent with theory. As population of a state increases, more land is developed. The effect of income and agricultural land value are both significant but inconsistent with some of the theory. The theory discussed in this paper suggests that as agricultural land value increases, it is less likely to be developed. The empirical results show a mostly positive effect. I also expected increased median household income to have strictly positive effect on land development, but only found a positive effect for lower income states, while higher income states saw increases in land development with increases in median household income. Both effects could be a result of applying theory built on the assumption of only one monocentric city rather than an entire state. More work could be done on the theory of the spatial urban structure of larger areas like states. This research provides an empirical foundation for the idea that these variables do not have the same effect on a larger scale as they do on a smaller city-wide scale.

Future research can address some limitations of this study in several ways. By adding auto-spatial correlation terms to the regression model, you can better control for the effects that neighboring states have on the land development of each state. It might also be possible to use a geographically weighted regression tool in a Geographic Information System (GIS) package such as ArcGIS to better account for the spatial effects of the regression (Xiao et al, 2006). Shape files of national land cover data are available from the National Land Cover Database that could be combined with other shapefiles containing economic information like gas prices to more closely analyze the spatial effects. It would also be helpful to add more variables to account for the other aspects of commuting costs such as opportunity cost of time.

WORKS CITED

BEA. (2018) *Gross Domestic Product by State Estimation Methodology*. Bureau of Economic Analysis, U.S. Department of Commerce.

BEA. (2018) *Annual Gross Domestic Product by State, GDP in Current Dollars.* Bureau of Economic Analysis, U.S. Department of Commerce. https://www.bea.gov/iTable/iTable.cfm?reqid=70&step=1&isuri=1&acrdn=2#reqid=70&step=1&isuri=1&7003=900&7004=sic&7005=-1&7006=00000&7001=1900&7002=1&7090=70&7093=levels

BEA. (2018) *Annual Gross Domestic Product by State, Quantity Indexes for Real GDP.* Bureau of Economic Analysis, U.S. Department of Commerce. https://www.bea.gov/iTable/iTable.cfm?reqid=70&step=1&isuri=1&acrdn=2#reqid=70&step=1&isuri=1&7003=900&7004=sic&7005=-1&7006=00000&7001=1900&7002=1&7090=70&7093=levels

Bounoua L., R. S. Defries, G. J. Collatz and P. J. Sellers, and H. Khan. (2002) Effects of Land Cover Conversions on Surface Climate, *Climatic Change* 52 (1-2): 29-64.

Brueckner, Jan K. (2000) Urban Sprawl: Diagnosis and Remedies, *International Regional Science Review* 23 (2): 160-171

Brueckner, Jan K. and David A. Fansler. (1983) The Economics of Urban Sprawl: Theory and Evidence on the Spatial Sizes of Cities, *The Review of Economics and Statistics* 65 (3): 479-482.

Burnett, Perry. (2012) Urban Industrial Composition and the Spatial Expansion of Cities, *Journal of Land Economics* 88 (4): 764-781

Census Bureau. (2016) *Historical Income Tables: Households, Median Household Income by State.* U.S. Census Bureau. https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-income-households.html

Chakir, Raja, Olivier Parent. (2009) Determinants of land use changes: A spatial multinomial probit approach, *Papers in Regional Science* 88 (2): 327-344.

EIA. (2018) State Energy Data System 1960-2016 (Complete) United States Energy Information Administration https://www.eia.gov/state/seds/seds-data-complete.php?sid=US#PricesExpenditures

Irwin, Elena G. and Nancy E. Bockstael. (2002) Interacting Agents, Spatial Externalities, and the Endogenous Evolution of Residential Land Use Patterns, *Journal of Economic Geography* 2 (1): 31-54.

Kahn, Matthew E. (2000) The Environmental Impact of Suburbanization, *Journal of Policy Analysis and Management* 19 (4): 569-586.

Kolko, Jed. (2017) Americans' Shift To The Suburbs Sped Up Last Year, *FiveThirtyEight,* www.fivethirtyeight.com

Liu, Yong, Wenze Yue, Peilei Fan. (2011) Spatial Determinants of Urban Land Conversion in Large Chinese Cities: A Case of Hangzhou, *Environment and Planning B: Urban Analytics and City Science* 38 (4): 706-725

Lubowski, Ruben N., Andrew J. Plantinga, and Robert N. Stavins. (2008) What Drives Land-Use Change in the United States? A National Analysis of Landowner Decisions, *Land Economics* 84 (4): 529-550.

McCann, Philip. (2001) *Urban and Regional Economics*. Oxford and New York:

Oxford University Press.

McGrath, Daniel T. (2005) More Evidence on the Spatial Scale of Cities, *Journal of Urban Economics* 58 (1): 1-10.

NASS. (1995) *Agricultural Land Values 1995.* National Agricultural Statistics Service, U.S. Department of Agriculture http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do;jsessionid=F154BA78C7C50C021C8CA924EDB72FD5?documentID=1446

NASS. (2005) *Agricultural Land Values 2005.* National Agricultural Statistics Service, U.S. Department of Agriculture http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do;jsessionid=F154BA78C7C50C021C8CA924EDB72FD5?documentID=1446

NASS. (2015) *Agricultural Land Values 2015.* National Agricultural Statistics Service, U.S. Department of Agriculture http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do;jsessionid=F154BA78C7C50C021C8CA924EDB72FD5?documentID=1446

NRCS. (2012) *Summary Report 2012 National Resources Inventory*. Natural Resources Conservation Service, U. S. Department of Agriculture in cooperation with Iowa State University Statistical Laboratory

Templeton, Scott R. and Sharma Ritu. Demographic, Economic, Environmental, and Political Determinants of the Spatial Scale of Statewide Developed Land in the U. S. Unpublished Manuscript.

Wang, Yiyi, Kara Krockelman, and Xiaokun Wang. (2011) Anticipation of Land Use

Change Through Use of Geographically Weighted Regression Models for Discrete Response, *Journal of the Transportation Research Board* 2245

Wheaton, William C. (1974) A Comparative Static Analysis of Urban Spatial Structure, *Journal of Economic Theory* 9 (2): 223-237.

Xiao, Jieying, Yanjun Shin, Jingfeng Ge, Ryutaro Tateishi, Changyuan Tang, Yanqing Liang, Zhiying Huang. (2006) Evaluating urban expansion and land use change in Shijiazhuang, China, by using GIS and remote sensing, *Landscape and urban Planning* 75 (1-2): 69-80.